University of Groningen

Hankel Singular Value Functions from Schmidt Pairs for Nonlinear Input-Output Systems

Gray, W. Steven; Scherpen, Jacquelien M.A.

Published in:
Proceedings of the 2002 American Control Conference

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2002

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Gray, W. S., & Scherpen, J. M. A. (2002). Hankel Singular Value Functions from Schmidt Pairs for Nonlinear Input-Output Systems. In *Proceedings of the 2002 American Control Conference* (Vol. 5, pp. 3540-3545). University of Groningen, Research Institute of Technology and Management.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 04-05-2023

Proceedings of the American Control Conference
Anchorage, AK May 8-10, 2002

Hankel Singular Value Functions from Schmidt Pairs for Nonlinear Input-Output Systems

W. Steven Gray

Department of Electrical
and Computer Engineering
Old Dominion University
Norfolk, Virginia 23529-0246
U.S.A.
gray@ece.odu.edu

Jacquelien M. A. Scherpen

Delft University of Technology
Fac. of Inf. Techn. & Syst.
Department of Electrical Engineering
P.O. Box 5031, 2600 GA Delft
The Netherlands
J.M.A.Scherpen@et.tudelft.nl

Abstract

In this paper three results are presented in singular value analysis of Hankel operators for nonlinear input-output systems. First the notion of a Schmidt pair is introduced which makes it possible to describe a Hankel singular value function from a purely input-output point of view. If a state space realization is known to exist then a set of sufficient conditions is provided for the existence of a Schmidt pair. Finally, it is shown that in a certain coordinate frame it is possible to relate this new singular value function definition to the existing definition due to Scherpen.

1 Introduction

Hankel theory for continuous-time nonlinear systems is considerably less developed than its linear counterpart. The classic result along these lines is due to Fliess [2, 3] who used a system Hankel matrix to describe when an analytic finite dimensional affine realization of an input-output system described by a certain functional series is minimal. This matrix in essence plays the same role that the system Hankel matrix does in linear and bilinear system theory [9, 10]. In a purely state space setting, the notion of Hankel singular values was generalized to nonlinear systems by Scherpen in [12, 13] and applied to model reduction problems. Connections between minimality and these invariants were later described in [15]. In [6, 14] a system Hankel operator was introduced for a general nonlinear input-output system and shown to be related, albeit in a fairly weak sense, to the singular value functions of Scherpen when the input-output operator had a finite dimensional state space realization.

In this paper three innovations are presented. First the notion of a Schmidt pair is introduced for a nonlinear input-output map. Using this notion it is then possible to describe in a coordinate free manner a Hankel singular value function from a purely input-output point of view. That is, as in the linear case, the notion of a Hankel singular value is described as an intrinsic property of the input-output system. However, if a state space realization is known to exist then a set of sufficient conditions is provided for the existence of a Schmidt pair. When these two results are combined, this provides a Hankel singular value function *representation* that is more directly tied to the Hankel operator than the existing notion is. Finally, it is shown that in a certain coordinate frame it is possible to relate this new singular value function definition to the existing definition of Scherpen. It is believed that this new definition may help solve a certain nonuniqueness property for nonlinear balanced realizations reported in [8, 16].

The paper is organized as follows. In Section 2, the nonlinear Hankel operator definition is reviewed in a more general context than it first appeared in [6, 14]. In Section 3 the notion of a Hilbert adjoint operator is briefly reviewed. This material is essential for understanding how to interpret the generalized Schmidt pair. The new results are all contained in Section 4. The final section summarizes the main conclusions of the paper and presents some remarks about open problems.

The mathematical notation used throughout is fairly standard. $\mathbb{R}^+$ denotes the set of nonnegative real numbers. The inner product and corresponding norm on $\mathbb{R}^n$ are represented, respectively, as $\langle x, y\rangle = x^T y$ and $\|x\| = \sqrt{\langle x, x\rangle}$. $L_p^i[a,b]$ represents the set of Lebesgue measurable functions, i-component vector-valued, with finite L_p norm, $\|\cdot\|_{L_p}$. The inner product on $L_2^i[a,b]$ is denoted by

$$\langle f, g\rangle_{L_2} = \int_a^b f(t)^T g(t)\, dt.$$

2 Hankel Operators Induced from Input-Output Systems

Let F be an input-output system defined on a set of admissible inputs $U[t_0,t_1]$ over the time interval $[t_0,t_1]$. The *time reversal* operator is taken as the injective mapping

$$\begin{aligned} \mathcal{R} \ &: \ U[t_0,t_1] \to U[-t_1,-t_0] \\ &: \ u \to \hat{u}(t) = u(-t), \end{aligned}$$

and the catenation of two signals $(u,v) \in U[t_0,t_1] \times U[t_2,t_3]$ at $\tau \in [t_0,t_1]$ is defined as

$$(u \#_\tau v)(t) = \begin{cases} u(t) & : \ t_0 \le t \le \tau \\ v((t-\tau)+t_2) & : \ \tau < t \le \tau + (t_3 - t_2). \end{cases}$$

It is generally assumed for any $\tau \in [t_0,t_1]$ that $U[t_0,t_1] = U[t_0,\tau]\#_\tau U[\tau,t_1]$.

0-7803-7298-0/02/$17.00 © 2002 AACC

Definition 2.1 *For any input-output system $F : U[t_0, t_1] \mapsto Y[t_0, t_1]$ with $t_0 < 0 < t_1$, the corresponding* **Hankel operator** *is*

$$\begin{aligned} \mathcal{H}_F \ &: \ U[0, -t_0] \times U[0, t_1] \to Y[0, t_1] \\ &: \ (u_-, u_+) \to y(t) = F(\underbrace{\mathcal{R}(u_-)\#_0 u_+}_{u})(t) \end{aligned}$$

for all $t \in [0, t_1]$.

The usual interpretation from linear system theory that $\mathcal{H}_F$ maps *past inputs* to *future outputs* is recovered from this definition when F is causal and homogeneous (i.e., $F(0) = 0$). In this context, the zero-input (for positive time) Hankel

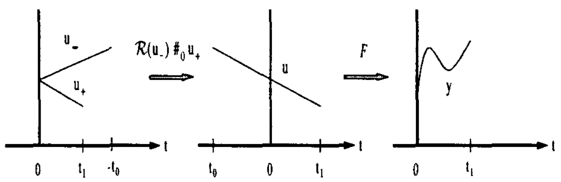

Figure 1: The Hankel operator, $\mathcal{H}_F$, corresponding to an input-output mapping F.

operator is denoted by $\mathcal{H}_{F,0}(\hat{u}) = \mathcal{H}_F(\hat{u}, 0)$.

Two inputs $u_-, v_- \in U[0, -t_0]$ are considered equivalent, i.e, $u_- \sim v_-$, when $\mathcal{H}_F(u_-, u_+) = \mathcal{H}_F(v_-, u_+)$ for every $u_+ \in U[0, t_1]$. Each equivalence class under this relation corresponds to the *state* of the system at time $t = 0$. When the quotient set $U[0, -t_0]/\sim$ is locally isomorphic to $\mathbb{R}^n$ then there corresponds an n dimensional state space realization of F. Our main interest is in operators that have affine input realizations

$$\begin{aligned} \dot{x} &= f(x) + g(x)u, \ \ x(t_0) = x_0 \\ y &= h(x) \end{aligned} \tag{1}$$

in terms of local coordinates on an n-dimensional state manifold $\mathcal{M}$. When F is homogeneous, it is always assumed that $f(0) = 0$ and $h(0) = 0$. The existence of any state space realization valid near $t = 0$ produces factorizations of $\mathcal{H}_F$ and $\mathcal{H}_{F,0}$ in terms of the controllability and observability operators

$$\begin{aligned} \mathcal{C} \ &: \ U[0, -t_0] \mapsto \mathcal{M} \\ &: \ u_- \mapsto \phi(t, t_0, x_0, \mathcal{R}(u_-))|_{t=0} \end{aligned}$$

and

$$\begin{aligned} \mathcal{O} \ &: \ U[0, t_1] \times \mathcal{M} \mapsto Y[0, t_1] \\ &: \ (u_+, x(0)) \mapsto h(\phi(t, 0, x(0), u_+)), \end{aligned}$$

respectively, where $\phi(t, t_0, x_0, u)$ denotes the solution of state equation (1) with $x(t_0) = x_0$ and any admissible input u applied. Specifically, then

$$\mathcal{H}_F(u_-, u_+) = \mathcal{O}(u_+, \mathcal{C}(u_-))$$

and

$$\mathcal{H}_{F,0}(\hat{u}) = \mathcal{O}(0, \mathcal{C}(\hat{u})) =: \mathcal{O}_0\mathcal{C}(\hat{u}).$$

Now to describe a Hankel singular value function in this setting the notion of an adjoint operator is needed.

3 Hilbert Adjoints of Nonlinear Operators

It is assumed throughout that F is L_2-stable in the sense that $u \in L_2^m(-\infty, 0]$ implies that $F(u)$ restricted to $[0, \infty)$ is in $L_2^\ell[0, \infty)$. In this case, the corresponding zero-input Hankel operator assumes the form $\mathcal{H}_{F,0} : L_2^m[0, \infty) \to L_2^\ell[0, \infty)$. Viewed as a mapping between Hilbert spaces, it is possible to compute a Hilbert adjoint of $\mathcal{H}_{F,0}$. Various nonlinear extensions of Hilbert adjoints have appeared in [1, 7, 11]. The following definition, which is fully developed in [7, 17] is most natural for the applications considered here.

Definition 3.1 *Given two Hilbert spaces H_1 and H_2, an operator $\mathcal{T} : H_1 \mapsto H_2$ has a global* ***nonlinear Hilbert adjoint*** *when there exists an operator $\mathcal{T}^* : H_2 \times H_1 \to H_1$ such that*

$$\langle \mathcal{T}(u), y \rangle_{H_2} = \langle u, \mathcal{T}^*(y, u) \rangle_{H_1}, \quad \forall u \in H_1, \quad \forall y \in H_2, \tag{2}$$

where $\mathcal{T}^*(y, u)$ is linear in y.

It is often the case that there exists a collection of nontrivial mappings (linear and nonlinear in y) of the form $\mathcal{B} : H_2 \times H_1 \mapsto H_1$ such that $\langle u, \mathcal{B}(y, u) \rangle_{H_1} = 0$, $\forall u \in H_1$, $\forall y \in H_2$. In which case, any adjoint mapping $\mathcal{T}^*$ is not uniquely defined since $\mathcal{T}^* + \mathcal{B}$ will also satisfy equation (2). In these circumstances, an adjoint operator should be viewed as a member of an equivalence class where two such operators $\mathcal{T}^*$ and $\mathcal{T}^{*'}$ are equivalent when

$$\langle u, \mathcal{T}^*(y, u) \rangle_{H_1} = \langle u, \mathcal{T}^{*'}(y, u) \rangle_{H_1}, \quad \forall u \in H_1, \quad \forall y \in H_2. \tag{3}$$

A shorthand notation for (3) is simply $\mathcal{T}^*(y, u) = \mathcal{T}^{*'}(y, u)$. Thus, any equality involving adjoint operators really means that both expressions belong to the same equivalence class. (See [8, 16] for analysis and examples closely related to this issue.) It is not necessary in many applications to have a globally defined $\mathcal{T}^*$. The following theorem provides a sufficient condition for the existence of a locally defined adjoint operator.

Theorem 3.1 *[7, 17] Suppose H_1 and H_2 are two Hilbert spaces and $U \subset H_1$ is any convex neighborhood of 0. Let $\mathcal{T} : U \mapsto H_2$ be a continuously Fréchet differentiable mapping on U such that $\mathcal{T}(0) = 0$. Then the mapping*

$$\mathcal{T}^*(y, u) = \int_0^1 (D\mathcal{T}(tu))^*(y)\, dt$$

is a suitable Hilbert adjoint of $\mathcal{T}$ on $H_2 \times U$, where $D\mathcal{T}$ is the Fréchet derivative of $\mathcal{T}$, and $(\cdot)^*$ denotes the usual linear adjoint operator.

While a useful device in many circumstances, a nonlinear Hilbert adjoint operator does not share all of the familiar properties associated with linear adjoints. For example, the sense in which operators can be composed when adjoint operators are present is more complicated since the domain of an adjoint operator is not simply the codomain of the original operator. For example, consider the Hilbert spaces H_i, $i = 1, 2, 3$, the operators

$$\begin{array}{ll} \mathcal{T}: H_1 \mapsto H_2 & \mathcal{S}: H_2 \mapsto H_3 \\ \quad : u \mapsto w & \quad : w \mapsto y \end{array}$$

and the corresponding adjoints

$$\begin{array}{ll} \mathcal{T}^*: H_2 \times H_1 \mapsto H_1 & \mathcal{S}^*: H_3 \times H_2 \mapsto H_2 \\ \quad : (w, u) \mapsto \bar{u} & \quad : (y, w) \mapsto \bar{w}. \end{array}$$

Clearly the composition and its adjoint

$$\begin{array}{ll} \mathcal{ST}: H_1 \mapsto H_3 & (\mathcal{ST})^*: H_3 \times H_1 \mapsto H_1 \\ \quad : u \mapsto y & \quad : (y, u) \mapsto \bar{u}. \end{array}$$

are well defined, but no *direct* composition like $\mathcal{T}^*\mathcal{T}$ or $\mathcal{T}^*\mathcal{S}^*$ is possible as in the classic setting. Still some *formal* compositions can be defined which have great utility in a variety of situations.

Definition 3.2 *Let H_i, $i = 1, 2, 3$, be a collection of Hilbert spaces. Assume $\mathcal{T} : H_1 \mapsto H_2$ and $\mathcal{S} : H_2 \mapsto H_3$ are two operators with well-defined adjoint operators. Define the following operator products:*

$$\begin{array}{rcll} (\mathcal{S}^*\mathcal{T})_1 & : & H_1 \times H_2 \mapsto H_2 & [\text{when } H_2 = H_3] \\ & : & (u, w) \mapsto \mathcal{S}^*(\mathcal{T}(u), w) & \end{array}$$

$$\begin{array}{rcl} (\mathcal{S}^*\mathcal{T})_2 & : & H_3 \times H_1 \mapsto H_1 \\ & : & (y, u) \mapsto \mathcal{S}^*(y, \mathcal{T}(u)). \end{array}$$

Of particular interest in the next section is the self-adjoint operator $\mathcal{H}^*_{F,0}\mathcal{H}_{F,0}(u) := (\mathcal{H}^*_{F,0}\mathcal{H}_{F,0})_1(u, u)$.

4 Hankel Singular Value Functions from Schmidt Pairs

We first develop the notion of a singular value function in a *coordinate free* setting. This is accomplished by defining a Schmidt pair for the operator $\mathcal{H}_{F,0}$. Let $\pi : L_2^m[0,\infty) \mapsto L_2^m[0,\infty)/\sim$ denote the canonical projection induced by $\mathcal{H}_{F,0}$. For any nonzero $\hat{v} \in L_2^m[0,\infty)$ and real numbers $0 < a < 1$, $b > 0$ define $V(a,b) = \{\hat{v}_\epsilon = \epsilon\,\hat{v} : \epsilon \in (1-a, 1+b)\}$. An operator $\mathcal{U}$ defined on $V(a,b)$ is called *homogeneous of degree one* when $\mathcal{U}(\hat{v}_\epsilon) = \epsilon\,\mathcal{U}(\hat{v})$.

Definition 4.1 *A **Schmidt pair** $(\hat{v}, \mathcal{U})$ for a Hankel operator $\mathcal{H}_{F,0} : L_2^m[0,\infty) \mapsto L_2^\ell[0,\infty]$ and some given adjoint operator $\mathcal{H}^*_{F,0}$ is a nonzero function $\hat{v} \in L_2^m[0,\infty)$ and an operator $\mathcal{U} : V(a,b) \mapsto L_2^\ell[0,\infty]$, homogeneous of degree one, such that*

$$\begin{aligned} \mathcal{H}_{F,0}(\hat{v}_\epsilon) &= \hat{\sigma}(\pi(\hat{v}_\epsilon))\,\mathcal{U}(\hat{v}_\epsilon) \\ \mathcal{H}^*_{F,0}(\mathcal{U}(\hat{v}_\epsilon), \hat{v}_\epsilon) &= \hat{\sigma}(\pi(\hat{v}_\epsilon))\,\hat{v}_\epsilon, \end{aligned}$$

for all $\hat{v}_\epsilon \in V(a,b)$ and some function $\hat{\sigma} : V(a,b)/\sim\, \mapsto \mathbb{R}^+$.

When such a pair $(\hat{v}, \mathcal{U})$ exists with $\hat{\sigma}$ and $\mathcal{U}(\hat{v})$ nonzero, $\mathcal{H}_{F,0}$ is locally injective on $V(a,b)$. That is, for any $\hat{v}_\epsilon, \hat{v}_{\epsilon'} \in V(a,b)$ such that $\mathcal{H}_{F,0}(\hat{v}_\epsilon) = \mathcal{H}_{F,0}(\hat{v}_{\epsilon'})$:

$$\begin{aligned} \pi(\hat{v}_\epsilon) = \pi(\hat{v}_{\epsilon'}) \;&\Rightarrow\; \mathcal{U}(\hat{v}_\epsilon) = \mathcal{U}(\hat{v}_{\epsilon'}) \\ &\Rightarrow\; \epsilon\,\mathcal{U}(\hat{v}) = \epsilon'\mathcal{U}(\hat{v}) \\ &\Rightarrow\; \epsilon = \epsilon' \\ &\Rightarrow\; \hat{v}_\epsilon = \hat{v}_{\epsilon'}. \end{aligned}$$

The linearity of $\mathcal{H}^*_{F,0}$ in its first argument implies directly that

$$\mathcal{H}^*_{F,0}(\mathcal{H}_{F,0}(\hat{v}_\epsilon), \hat{v}_\epsilon) = \hat{\sigma}^2(\pi(\hat{v}_\epsilon))\,\hat{v}_\epsilon,$$

so $\hat{\sigma}$ is logically called a *singular value function* for the operator pair $(\mathcal{H}_{F,0}, \mathcal{H}^*_{F,0})$. Singular values functions are strongly dependent on the choice of adjoint operators. For example, if $\mathcal{H}^{*'}_{F,0}$ is a second adjoint operator distinct from $\mathcal{H}^*_{F,0}$ then

$$\begin{aligned} \mathcal{H}^{*'}_{F,0}(\mathcal{H}_{F,0}(\hat{v}_\epsilon), \hat{v}_\epsilon) \;=\; & \hat{\sigma}^2(\pi(\hat{v}_\epsilon))\,\hat{v}_\epsilon + \\ & \hat{\sigma}(\pi(\hat{v}_\epsilon))\,\mathcal{B}(\mathcal{H}_{F,0}(\hat{v}_\epsilon), \hat{v}_\epsilon) \end{aligned}$$

for some function $\mathcal{B} : L_2^\ell[0,\infty) \times L_2^m[0,\infty) \mapsto L_2^m[0,\infty)$ with $\langle \hat{v}_\epsilon, \mathcal{B}(\mathcal{H}_{F,0}(\hat{v}_\epsilon), \hat{v}_\epsilon)\rangle = 0$ everywhere on $V(a,b)$. So clearly $\hat{\sigma}$ is not a singular value function for the pair $(\mathcal{H}_{F,0}, \mathcal{H}^{*'}_{F,0})$. Different adjoint operators can potentially produce different singular value functions. But they all share the property that

$$\left\langle \hat{v}_\epsilon, \mathcal{H}^*_{F,0}(\mathcal{H}_{F,0}(\hat{v}_\epsilon), \hat{v}_\epsilon)\right\rangle_{L_2} = \hat{\sigma}^2(\pi(\hat{v}_\epsilon))\,\langle \hat{v}_\epsilon, \hat{v}_\epsilon\rangle_{L_2}.$$

For a linear operator, $V(a,b)$ is normally taken as the span of $\hat{v}$ over $\mathbb{R}$ with $\|\hat{v}\| = 1$. For compact linear operators, constant singular values functions and linear $\mathcal{U}$ operators are known to always exist. In fact, the operator $\mathcal{H}_{F,0}$ has a singular value decomposition of the form

$$\mathcal{H}_{F,0}(\hat{u}) = \sum_{i=1}^{\infty} \hat{\sigma}_i\,\mathcal{U}_i(\hat{u})\,\langle \hat{v}_i, \hat{u}\rangle_{L_2}, \quad \forall \hat{u} \in L_2^m[0,\infty),$$

where each $(\hat{v}_i, \mathcal{U}_i)$ is a Schmidt pair, $\hat{\sigma}_i \geq \hat{\sigma}_{i+1}$ for $i \geq 1$, and $\{\hat{v}_i\}_{i=1}^\infty$ is a complete orthonormal set for $L_2^m[0,\infty)$. In the nonlinear setting, when a family of Schmidt pairs $\{(\hat{v}_i, \mathcal{U}_i)\}_{i=1}^\infty$ is known to exist, the analogous expression is

$$\mathcal{H}_{F,0}(\hat{u}) = \sum_{i=1}^{\infty} \hat{\sigma}_i(\pi(\hat{u}))\,\mathcal{U}_i(\hat{u})\,\langle \hat{v}_i, \hat{u}\rangle_{L_2}, \quad \forall \hat{u} \in V,$$

where $V := \bigcup_{i=1}^\infty V_i(a_i, b_i)$. But this expression is not necessarily valid anywhere else in $L_2^m[0,\infty)$. Also unlike the linear case, this decomposition will be highly nonunique when the set of adjoint operators for $\mathcal{H}_{F,0}$ is large. Thus, distinct decompositions truncated to the same number of leading terms will result in different approximations of $\mathcal{H}_{F,0}$. This has obvious consequences for any nonlinear model reduction algorithm based on singular values functions (see [8] for a related discussion).

When F is homogeneous with a smooth n dimensional state space realization $(f, g, h, 0)$, which is L_2 input-to-state stable on a neighborhood W of 0 (which means that when $u \in L_2^m(-\infty, 0]$, the corresponding state vector, $x(t)$, assuming the initial condition $x(-\infty) = 0$, is finite on $(-\infty, 0]$ and always contained in W), it is possible to prove the existence of n Schmidt pairs and singular value functions for $\mathcal{H}_{F,0}$. The state space context also provides a convenient representation for these functions. This is accomplished using the *energy functions* for $(f, g, h, 0)$ as described below.

Definition 4.2 *The* **controllability** *and* **observability functions** *for the system $(f, g, h, 0)$ are defined, respectively, as*

$$L_c(x) = \min_{\substack{u \in L_2(-\infty,0) \\ x(-\infty)=0,\ x(0)=x}} \frac{1}{2}\int_{-\infty}^{0} \|u(t)\|^2\, dt$$

and

$$L_o(x) = \frac{1}{2}\int_{0}^{\infty} \|y(t)\|^2\, dt,$$

when $x(0) = x$, and $u(t) = 0$ for $0 \le t < \infty$.

The following result is known.

Theorem 4.1 *[12] Consider a system $(f, g, h, 0)$ where*

(A1) *f is asymptotically stable on some neighborhood Y of 0;*

(A2) *The system (f, g, h) is zero-state observable on Y (i.e., $\mathcal{O}_0(x_0) \equiv 0$ implies that $x_0 = 0$);*

(A3) *L_c and L_o exist and are smooth on Y.*

There exists a coordinate transformation $x = \psi(z)$, $\psi(0) = 0$, defined on a neighborhood U of 0 which converts the system into an **input-normal/output-diagonal realization**, *where*

$$\begin{aligned}\tilde{L}_c(z) &:= L_c(\psi(z)) = \frac{1}{2}z^T z,\\ \tilde{L}_o(z) &:= L_o(\psi(z)) = \frac{1}{2}z^T diag(\tau_1(z), \ldots, \tau_n(z))z\end{aligned}$$

with $\tau_1(z) \ge \ldots \ge \tau_n(z)$ being smooth functions on $W := \psi^{-1}(U)$ provided the number of distinct $\tau_i(z)$'s are constant over W.

The set of functions τ_i, $i = 1, \ldots, n$ are called *singular value functions* of $(f, g, h, 0)$ in [12]. They should not be confused with singular value functions, $\hat{\sigma}_i$, for $(\mathcal{H}_{F,0}, \mathcal{H}_{F,0}^*)$, though as will be shown momentarily, there is a relationship between the two concepts. When $\tilde{L}_o$ is not in a diagonal form, the realization is said to simply be in *input-normal form*. It is also known that there exists a coordinate transformation $z = \eta(\bar{z})$, $\eta(0) = 0$, defined on the neighborhood W of 0 which converts the system into a *balanced realization*, where

$$\begin{aligned}\breve{L}_c(\bar{z}) &:= \tilde{L}_c(\eta(\bar{z}))\\ &= \frac{1}{2}\bar{z}^T diag(\hat{\tau}(\bar{z}_1)^{-1}, \ldots, \hat{\tau}(\bar{z}_n)^{-1})\bar{z}\\ \breve{L}_o(\bar{z}) &:= \tilde{L}_o(\eta(\bar{z}))\\ &= \frac{1}{2}\bar{z}^T diag(\hat{\tau}_1(\bar{z}_1)^{-1}\tau_1(\eta^{-1}(\bar{z})), \ldots,\\ &\qquad \hat{\tau}_n(\bar{z}_n)^{-1}\tau_n(\eta^{-1}(\bar{z})))\bar{z},\end{aligned}$$

with $\hat{\tau}_i(\bar{z}_i) := \tau_i(0, \ldots, 0, \eta_i^{-1}(\bar{z}_i), 0, \ldots, 0)^{\frac{1}{2}}$ for $i = 1, \ldots, n$. Along coordinate axes it is easily verified that

$$\begin{aligned}\breve{L}_c(0, \ldots, 0, \bar{z}_i, 0, \ldots, 0) &= \tfrac{1}{2}\bar{z}_i^2\hat{\tau}_i(\bar{z}_i)^{-1}\\ \breve{L}_o(0, \ldots, 0, \bar{z}_i, 0, \ldots, 0) &= \tfrac{1}{2}\bar{z}_i^2\hat{\tau}_i(\bar{z}_i).\end{aligned}$$

To relate the singular values functions of $(\mathcal{H}_{F,0}, \mathcal{H}_{F,0}^*)$ to those of a given state space realization of F, the key idea is to select the adjoint operator for $\mathcal{H}_{F,0}$ in a manner *consistent* with the realization. This is done in the following two theorems. The first theorem is adapted from [4, 5]. It expresses an adjoint operator in terms of a solution to a two-point boundary value problem with conjugate points $t_0 = -\infty$ and $t_1 = \infty$. Its original proof is done by viewing the system as a port-controlled Hamiltonian system. The second theorem applies the first theorem. It provides sufficient conditions for the existence of a Schmidt pair using the particular adjoint operator described below.

Theorem 4.2 *[4, 5] Let F be a causal homogeneous L_2-stable input-output mapping with a smooth n dimensional state space realization $\Sigma = (f, g, h, 0)$ that is L_2 input-to-state stable on a neighborhood W of 0. Consider the mapping*

$$\begin{aligned}\mathcal{H}_{\Sigma,0}^* &: L_2^\ell[0,\infty) \times L_2^m[0,\infty) \mapsto L_2^m[0,\infty)\\ &: (u_a, \hat{u}) \mapsto y_a\end{aligned}$$

defined by the state space realization

$$\dot{x} = f(x) + g(x)\mathcal{R}_-(\hat{u}), \quad x(-\infty) = 0 \tag{4}$$

$$\dot{p} = -A^T(x)p - C^T(x)u_a, \quad p(\infty) = 0 \tag{5}$$

$$y_a = \mathcal{R}_+(g^T(x)p) \tag{6}$$

with

$$\mathcal{R}_+(u) = \begin{cases} u(-t) & : t \in [0,\infty) \\ 0 & : t \in (-\infty, 0] \end{cases}$$

$$\mathcal{R}_-(\hat{u}) = \begin{cases} 0 & : t \in [\infty, 0] \\ \hat{u}(-t) & : t \in (-\infty, 0] \end{cases},$$

$A(x) \in \mathbb{R}^{n\times n}$, $C(x) \in \mathbb{R}^{\ell\times n}$ such that $f(x) = A(x)x$ and $h(x) = C(x)x$ for all $x \in W$, and where it is assumed a priori that $p(-\infty)$ is finite for all $(u_a, \hat{u}) \in L_2^\ell[0,\infty) \times L_2^m[0,\infty)$. (By assumption, $u_a(t) = 0$ when $t \le 0$.) Then $H_{\Sigma,0}^*$ is a valid Hilbert adjoint of $\mathcal{H}_{F,0}$. That is,

$$\langle \mathcal{H}_{F,0}(\hat{u}), u_a\rangle_{L_2} = \langle \hat{u}, \mathcal{H}_{\Sigma,0}^*(u_a, \hat{u})\rangle_{L_2},$$

for all $u_a \in L_2^\ell[0,\infty)$ and $\hat{u} \in L_2^m[0,\infty)$.

It should be noted that the factorizations $f(x) = A(x)x$ and $h(x) = C(x)x$ are always possible since it is assumed that $f(0) = 0$ and $h(0) = 0$. But, as is well known, these factorizations are not unique. This means that potentially a *set* of consistent adjoints for $\mathcal{H}_{F,0}$ is possible. The L_2 input-to-state stability of Σ and the assumption that $p(-\infty)$ is finite insures that the state space realization for $\mathcal{H}_{\Sigma,0}^*$ is well defined solution for all time and every admissible input.

Theorem 4.3 *Let $\Sigma = (f, g, h, 0)$ be a smooth L_2 input-to-state stable realization in input-normal form of a causal L_2-stable input-output mapping F on a neighborhood W of 0. Let $\tilde{x}$ be a fixed nonzero state in W, $v^\dagger$ the minimum energy input which drives $x(-\infty) = 0$ to $x(0) = \tilde{x}$ and $\hat{v}^\dagger = \mathcal{R}_+(v^\dagger)$. Suppose the following assumptions are valid:*

(B1) *There exists a factorization $f(x) = A(x)x$ where the symmetric part $A_s(x) := (A(x)+ A^T(x))/2 = -g^T(x)g(x)/2$ for all $x \in W$.*

(B2) *The system (4)-(5) with $(u_a, \hat{u}) = (\mathcal{H}_{F,0}(\hat{v}^\dagger), \hat{v}^\dagger)$ has a well defined solution with the property that $p(0) = \hat{\sigma}^2(\tilde{x})\tilde{x}$ for some positive constant $\hat{\sigma}^2(\tilde{x})$.*

Then the operator pair $(\mathcal{H}_{F,0}, \mathcal{H}^*_{\Sigma,0})$ has at least one Schmidt pair.

Proof: The proof is constructive. Since Σ is in input-normal form, there exists sufficiently small constants $0 < a < 1$ and $b > 0$ such that $V(a,b)$ corresponds to the set of minimum energy inputs, $v^\dagger_\epsilon$, which drive the state $x(-\infty) = 0$ to $x(0) = \tilde{x}_\epsilon := \epsilon\tilde{x}$ for any $\epsilon \in (1-a, 1+b)$. Now define the *function* $\hat{\sigma}$ by assigning $\hat{\sigma}(\tilde{x}_\epsilon) = \hat{\sigma}(\tilde{x})/\epsilon$ for any $\epsilon \in (1-a, 1+b)$. This in turn provides a well defined homogeneous operator of degree one

$$\begin{aligned} \mathcal{U} &: V(a,b) \mapsto L_2^\ell[0,\infty) \\ &: \hat{v}^\dagger_\epsilon \mapsto \mathcal{H}_{F,0}(\hat{v}^\dagger_\epsilon)/\hat{\sigma}(\tilde{x}_\epsilon) = \epsilon\mathcal{H}_{F,0}(\hat{v}^\dagger_\epsilon)/\hat{\sigma}(\tilde{x}). \end{aligned}$$

The claim is that $(\hat{v}^\dagger, \mathcal{U})$ is a Schmidt pair for $(\mathcal{H}_{F,0}, \mathcal{H}^*_{\Sigma,0})$. By design it is clear that $\mathcal{H}_{F,0}(\hat{v}^\dagger_\epsilon) = \hat{\sigma}(\tilde{x}_\epsilon)\mathcal{U}(\hat{v}^\dagger_\epsilon)$ for all $\hat{v}^\dagger_\epsilon \in V(a,b)$. One can view $\tilde{x}_\epsilon = \pi(\hat{v}^\dagger_\epsilon)$ such that $\sigma = \hat{\sigma} \circ \pi$ is a singular value function of $\mathcal{H}_{F,0}$ in the sense of Definition 4.1. To verify the rest of this definition, first recall that in [12] it was shown that L_c is the smooth solution of the partial differential equation

$$\frac{\partial L_c}{\partial x} f(x) + \frac{1}{2}\frac{\partial L_c}{\partial x} g(x)g^T(x)\frac{\partial L_c^T}{\partial x} = 0. \qquad (7)$$

Furthermore, $v^\dagger = g^T(x)\partial L_c^T/\partial x(x)$ when evaluated along the solution of $\dot{x} = f(x) + g(x)v^\dagger$ starting at $x(-\infty) = 0$ and terminating at $x(0) = \tilde{x}$. In input-normal form this implies that $v^\dagger = g^T(x)x$. Now setting $u_a = \mathcal{H}_{F,0}(\hat{v}^\dagger)$ and $\hat{u} = \hat{v}^\dagger$, the realization of $\mathcal{H}^*_{\Sigma,0}$ evaluated at these inputs has the equivalent form for $t \leq 0$:

$$\begin{aligned} \dot{x} &= f(x) + g(x)g^T(x)x, \;\; x(0) = \tilde{x} \qquad (8) \\ \dot{p} &= -A^T(x)p, \;\; p(0) = \tilde{p}, \qquad (9) \end{aligned}$$

for some $\tilde{p} \in \mathbb{R}^n$. Equation (7) reduces to

$$x^T \left[A(x) + A^T(x) + g(x)g^T(x)\right] x = 0.$$

This does *not* in general imply that $A(x) + A^T(x) + g(x)g^T(x) = 0$, but if this is the case, i.e., if assumption (B1) is satisfied, then equations (8)-(9) become

$$\begin{aligned} \dot{x} &= (A(x) + g(x)g^T(x))x, \;\; x(0) = \tilde{x} \qquad (10) \\ \dot{p} &= (A(x) + g(x)g^T(x))p, \;\; p(0) = \tilde{p}. \qquad (11) \end{aligned}$$

Now if $\tilde{p}$ and $\tilde{x}$ are related by the constant $\hat{\sigma}^2(\tilde{x})$, assumption (B2), then equations (10) and (11) will have solutions that are related by this same constant for all negative time. That is, $p(t) = \hat{\sigma}^2(\tilde{x})x(t)$ when $t \leq 0$. Hence,

$$\begin{aligned} \mathcal{H}^*_{\Sigma,0}(\mathcal{H}_{F,0}(\hat{v}^\dagger), \hat{v}^\dagger) &= \mathcal{R}_+(g^T(x)p) \\ &= \mathcal{R}_+(g^T(x)\hat{\sigma}^2(\tilde{x})x) \\ &= \hat{\sigma}^2(\tilde{x})\mathcal{R}_+(g^T(x)x) \\ &= \hat{\sigma}^2(\tilde{x})\mathcal{R}_+(v^\dagger) \\ &= \hat{\sigma}^2(\tilde{x})\hat{v}^\dagger. \end{aligned}$$

Using the linearity of the first argument of $\mathcal{H}^*_{\Sigma,0}$, it then follows that

$$\mathcal{H}^*_{\Sigma,0}(\mathcal{U}(\hat{v}^\dagger), \hat{v}^\dagger) = \hat{\sigma}(\tilde{x})\hat{v}^\dagger.$$

Multiplying both sides of the equation by ϵ gives

$$\mathcal{H}^*_{\Sigma,0}(\mathcal{U}(\hat{v}^\dagger_\epsilon), \hat{v}^\dagger) = \hat{\sigma}(\tilde{x})\hat{v}^\dagger_\epsilon,$$

and observe that setting instead $\hat{u} = \hat{v}^\dagger_\epsilon$ above corresponds to replacing $\tilde{x}$ with $\tilde{x}_\epsilon$ to produce the desired result

$$\mathcal{H}^*_{\Sigma,0}(\mathcal{U}(\hat{v}^\dagger_\epsilon), \hat{v}^\dagger_\epsilon) = \hat{\sigma}(\tilde{x}_\epsilon)\hat{v}^\dagger_\epsilon, \;\; \hat{v}^\dagger_\epsilon \in V(a,b).$$

■

The factorization property in (B1) is automatically satisfied in the linear setting because for any input normal form (A, B) the Lyapunov equation $A + A^T + BB^T = 0$ is satisfied. But in the nonlinear case not much is known about these types of factorizations. (A related factorization is described and characterized in [8].) Fortunately, the boundary property in (B2) can be assured when the realization is in input-normal/output-diagonal form. The result given below shows that the singular value functions defined for a Schmidt pair will coincide with the singular value functions defined in Theorem 4.1 when each is evaluated along a coordinate axis, and the vector $\tilde{x}$ in Theorem 4.3 is suitably chosen.

Theorem 4.4 *Let $\Sigma = (f, g, h, 0)$ be a smooth L_2 input-to-state stable realization in input-normal/ output-diagonal form of a causal homogeneous L_2-stable input-output mapping F in a neighborhood W of 0. Assume (A1)-(A3) are satisfied in Theorem 4.1, and let τ_i, $i = 1, \ldots, n$ be the corresponding singular value functions of Σ. Furthermore, let $\tilde{x}_i = (0, \ldots 0, x_i, 0 \ldots, 0)^T \in W$ for $i \in \{1, 2, \ldots, n\}$ with $x_i \neq 0$ and $\hat{v}_i$ the respective minimum energy input. If (B1) in Theorem 4.3 holds and the system (4)-(5) with $(u_a, \hat{u}) = (\mathcal{H}_{F,0}(\hat{v}^\dagger_i), \hat{v}^\dagger_i)$ has a well defined solution then:*

i. *The operator pair$(\mathcal{H}_{F,0}, \mathcal{H}^*_{\Sigma,0})$ has at least one Schmidt pair with singular value function $\hat{\sigma}_i$.*

ii. *It follows that $\hat{\sigma}_i(\tilde{x}_i) = \tau_i(\tilde{x}_i)$ for $i = 1, \ldots, n$.*

Proof: In the context of Theorem 4.3 let $(u_a, \hat{u}) = (\mathcal{H}_{F,0}(\hat{v}^\dagger_i), \hat{v}^\dagger_i)$. Then equations (4)-(5) become for $t > 0$:

$$\begin{aligned} \dot{x} &= f(x), \;\; x(0) = \tilde{x}_i \\ \dot{p} &= -A^T(x)p - c^T(x)h(x), \;\; p(0) \text{ given.} \end{aligned}$$

For any $x \in W$ it then follows that

$$x^T \dot{p} = -f^T(x)p - h^T(x)h(x), \ \ p(0) \text{ given},$$

or equivalently,

$$\frac{d(x^T p)}{dt} = -h^T(x)h(x).$$

Integrating both sides of the equation over the trajectory of $\{x(t), p(t)\}$ from $t = 0$ to $t = \infty$ produces

$$\begin{aligned} \frac{1}{2}\tilde{x}_i^T p(0) &= L_o(\tilde{x}_i) \\ &= \frac{1}{2}\tau_i(\tilde{x}_i)x_i^2. \end{aligned}$$

Since $x_i \neq 0$ the resulting boundary condition $p(0) = \tau_i(\tilde{x}_i)\tilde{x}_i$ assures that the operator pair has a well defined Schmidt pair, and thus has a singular value function $\hat{\sigma}_i$ in this sense. By the definition in Theorem 4.3 it follows immediately then that $\hat{\sigma}_i(\tilde{x}_i) = \tau_i(\tilde{x}_i)$. ■

5 Conclusions and Future Research

In this paper three results were presented. First the notion of a Schmidt pair was introduced for a nonlinear input-output map. Using this device it was then possible to describe a Hankel singular value as an intrinsic property of the input-output system. When a state space realization is known to exist, a set of sufficient conditions was provided for the existence of Schmidt pair. Finally, it was shown that in a certain coordinate frame this new singular value function definition can be related to the existing notion, which is defined entirely in term of normal form coordinates. An application of these results may be to provide a canonical structure for a balanced realization of a nonlinear system. This would in turn put the theory of nonlinear state space model reduction on firmer theoretical ground.

References

[1] V. Burýsková, *Some Properties of Nonlinear Adjoint Operators*, Rocky Mountain J. of Math., 28 (1998), pp. 41-59.

[2] M. Fliess, *Matrices de Hankel*, J. Math. Pures. Appl., 53 (1974), pp. 197-224.

[3] ——, *Réalisation Locale des Systèmes Non Linéaires, Algèbres de Lie Filtrées Transitives et Séries Génératrices Non Commutatives*, Invent. Math., 71 (1983), pp. 521-537.

[4] K. Fujimoto, J. M. A. Scherpen and W. S. Gray, *Hamiltonian Realizations of Nonlinear Adjoint Operators*, Proc. IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control, Princeton, NJ, 2000, pp. 39-44.

[5] ——, *Hamiltonian Realizations of Nonlinear Adjoint Operators*, Automatica, to appear.

[6] W. S. Gray and J. M. A. Scherpen, *Hankel Operators and Gramians for Nonlinear Systems*, Proc. IEEE 37th Conf. Decision and Control, Tampa, FL, 1998, pp. 1416-1421.

[7] ——, *Nonlinear Hilbert Adjoints: Properties and Applications to Hankel Singular Value Analysis*, Proc. 2001 American Control Conf., Arlington, VA, 2001, pp. 3582-3587.

[8] ——, *On the Nonuniqueness of Singular Value Functions and Balanced Nonlinear Realizations*, Systems & Control Letters, 44 (2001), pp. 219-232.

[9] A. Isidori and A. Ruberti, *Realization Theory of Bilinear Systems*, in Geometric Methods in System Theory, D. Q. Mayne and R. W. Brockett, eds., D. Reidel, Dordrecht, 1973.

[10] R. E. Kalman, P. L. Falb and M. A. Arbib, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.

[11] G. I. Marchuk, V. I. Agoshkov and V. P. Shutyaev, *Adjoint Equations and Perturbation Algorithms in Nonlinear Problems*, CRC Press, Boca Raton, Florida, 1996.

[12] J. M. A. Scherpen, *Balancing for Nonlinear Systems*, Systems & Control Letters, 21 (1993), pp. 143-153.

[13] ——, *On the Similarity Invariance of Balancing for Nonlinear Systems*, Preprints IFAC Nonlinear Control Systems Design Symp. (NOLCOS), Tahoe City, CA, 1995, pp. 783-788.

[14] J. M. A. Scherpen and W. S. Gray, *On Singular Value Functions and Hankel Operators for Nonlinear Systems*, Proc. 1999 American Control Conf., San Diego, CA, 1999, pp. 2360-2364.

[15] ——, *Minimality and Local State Decompositions of a Nonlinear State Space Realization Using Energy Functions*, IEEE Trans. Automatic Control, AC-45 (2000), pp. 2079-2086.

[16] ——, *State Dependent Matrices and Balanced Energy Functions for Nonlinear Systems*, Proc. 2000 Inter. Symp. Mathematical Theory of Networks and Systems (MTNS), Perpignan, France, 2000.

[17] ——, *Nonlinear Hilbert Adjoints: Properties and Applications to Hankel Singular Value Analysis*, Nonlinear Analysis: Theory, Methods & Applications, to appear.